УДК 681.3

О.В. Дерев'янченко
Київський національний університет імені Тараса Шевченка, Україна
alexanderder@mail.ru

Аналіз інформації у комп'ютерній мережі за допомогою системи «LingvoSniffer»

В статті розглядаються принципи побудови системи аналізу інформації у комп'ютерній мережі із застосуванням лінгвістичного модуля, яка може вирішувати актуальне питання контролю мережі як засобу передачі та обміну даними між користувачами та доступу до певної інформації.

Вступ

З розвитком комп'ютерних мереж все більш актуальним постає питання контролю мережі як засобу передачі та обміну даними між її користувачами. Це обумовлено необхідністю підтримання характеристик мережі на рівні, необхідному для забезпечення її дієздатності та виконання поставлених перед нею завдань в умовах сучасних політик безпеки в підприємницьких чи навіть шкільних мережах.

В статті розглядається система «LingvoSniffer», що призначена для аналізу та фільтрації мережних даних з метою подолання проблеми обміну інформацією, пов'язаною з недозволеним вмістом даних, що передаються. Тобто заборона передачі та прийому інформації, яка може містити деяку конфіденційну інформацію або заборонену для цільової аудиторії (наприклад, «доросла інформація», до якої можуть мати доступ неповнолітні).

З поширенням сфери використання комп'ютерних мереж перед користувачами та адміністраторами постають завдання, що можуть бути вирішені тільки за допомогою спеціалізованого програмного забезпечення [1], [2].

Серед таких завдань:

– аналіз інформації, якою обмінюються користувачі мережі, з метою збору статистичних даних;

– розподіл доступу до мережних ресурсів користувачам різних категорій та інші.

Існує окремий клас програмного забезпечення, що дозволяє вирішувати поставлені завдання. Це – аналізатори мережних пакетів, або сніфери та мережні фільтри (фаєрволи).

Метою даної статті є показати принципи побудови подібних систем захисту і контролю комп'ютерних мереж.

Побудова системи «LingvoSniffer»

Програмний пакет аналізу та захисту мережі складається з модулів, що дозволяють вирішувати окремі класи задач:

– сніфер – для збору та відображення мережного трафіка з метою подальшого аналізу;

– фаєрвол – фільтрація мережних пакетів за заданими правилами;

– лінгвістичний модуль – реалізує додаткові можливості роботи з лінгвістичною інформацією.

Система прослуховування мережного трафіка структурно повинна складатись як мінімум з двох частин:
– програмного драйвера, який забезпечує ефективну роботу з мережним адаптером комп'ютера на низькому апаратному рівні;
– клієнтської частини, яка забезпечує отримання даних від драйвера, декодування їх, відображення у потрібному форматі та діалог з користувачем системи.

Розглянемо докладніше методики побудови та принципи роботи кожного модуля.

Модуль сніфера

Сніфер – це програма, що дозволяє перехоплювати мережний трафік. Коли говорять про сніфери, то звичайно проводять аналогію з прослуховуванням телефонних розмов. Підключившись до телефонної мережі, можна перехопити спілкування людей. Приблизно так у комп'ютерних мережах можна перехоплювати інформацію, якою обмінюються комп'ютери. Прослуховування можливе завдяки особливості архітектури мережної технології Ethernet. Архітектура більшості локальних мереж заснована на технології Ethernet (ether – ефір, network – мережа), у якій усі пристрої підключені до одного середовища передачі даних і спільно його використовують. Використовуючи цю особливість Ethernet, відпадає необхідність несанкціонованого підключення до сегмента мережі – комп'ютер, з якого передбачається прослуховування, вже підключений до деякого сегмента мережі.

Сніфер може аналізувати тільки те, що проходить через його мережну карту. Всередині одного сегмента мережі Ethernet усі пакети розсилаються всім машинам, через це можливе перехоплення чужої інформації. Використання комутаторів (switch, hub) та їхня грамотна конфігурація вже є захистом від прослуховування.

Між сегментами інформація передається через комутатори. Комутація пакетів – форма передачі, при якій дані, розбиті на окремі пакети, можуть пересилатися з вихідного пункту в пункт призначення різними маршрутами. Так, якщо хтось в іншому сегменті посилає внутрішні пакети, то у ваш сегмент комутатор ці дані не відправить.

Програми, що прослуховують, чи пакетні аналізатори відносяться до класу утиліт подвійного призначення. З одного боку, сніфери – могутня зброя, за допомогою якої можна здійснити пасивну мережну атаку. Ці програми можуть являти собою серйозну загрозу, оскільки можуть перехоплювати і розшифровувати імена і паролі користувачів, конфіденційну інформацію, порушувати роботу окремих комп'ютерів і мережі в цілому. Відомо, що в більшості протоколів передачі даних (FTP, POP, HTTP та ін.) [3], [4] секретна інформація між клієнтом і сервером передається відкритим текстом. Тому зловмиснику не складає великої праці одержати доступ до чужої інформації. З іншого боку, сніфери допомагають системним адміністраторам здійснювати діагностику мережі і відслідковувати атаки комп'ютерних злочинців. Крім того, вони служать для перевірки і детального аналізу правильності конфігурації мережного програмного забезпечення.

Інформація передається по мережі, і одержує її кожен пристрій цієї мережі. За замовчуванням мережна плата комп'ютера бачить тільки те, що призначено саме для неї. Однак програми, що прослуховують, встановлюють її в режим прийому всіх пакетів – promiscuous mode. В основі багатьох сніферів були і є мережні драйвери і бібліотеки (libpcap, libnet). Для переключення мережної плати в promiscuous mode потрібно низькорівневе програмування її портів. У багатозадачній ОС таку роботу можуть виконати тільки драйвери рівня ядра системи (kernel-mode drivers). Перші програми такого типу були створені для операційних систем Unix.

Незабаром сніфери були реалізовані і в популярній ОС Windows, але їхня робота в цій системі також вимагала мережного драйвера, що переключав мережну плату (NIC – network interface card) у спеціальний режим [5].

У системі, що розглядається, програмна реалізація заснована на використанні бібліотеки winpcap [6] – аналога libpcap [7], але орієнтованої для роботи з платформою win32.

Інформація в мережі передається у вигляді пакетів даних спеціального формату – фреймів (або кадрів). Для того щоб мати можливість правильно аналізувати перехоплені дані, потрібно мати ґрунтовні знання про будову мережних пакетів різного призначення.

Модуль фаєрволу

Якщо сніфер призначено переважно для збору та аналізу мережної інформації, то призначення такого класу програм як фаєрволи полягає в іншому.

Фаєрвол (від англ. Firewall – вогняна стіна) – це один з програмно-апаратних методів захисту від мережних атак.

Взагалі до програмно-апаратних засобів захисту мережі відносять:

– апаратні шифратори мережного трафіка;
– захищені мережні протоколи;
– програмно-апаратні аналізатори мережного трафіка;
– захищені мережні ОС.

Модуль фаєрволу, що реалізовано, має можливість керування та налаштування за допомогою команд, які може викликати зовнішня програма.

Фільтрація мережного трафіка є основною функцією систем Firewall та дозволяє адміністратору безпеки комп'ютерної мережі централізовано проводити необхідну мережну політику безпеки у виділеному сегменті IP-мережі. Тобто, налаштувавши фаєрвол необхідним чином, можна дозволити або заборонити користувачам як доступ з зовнішньої мережі до відповідних служб хостів, або до хостів, що знаходяться в захищеному сегменті, так і доступ користувачам з внутрішньої мережі до відповідних ресурсів зовнішньої мережі (IP-адреси).

Лінгвістичний модуль

Основною задачею системи є збір лінгвістичної статистики. Система аналізує вміст мережних пакетів на наявність ключових слів та збирає статистику кількості цих слів, які пройшли по вказаному каналу. Реалізовано додаткову можливість генерування списку слів, які близькі до еталонного слова. Цей список слів шукається у базі знань WordNet [8]. Близькість визначається семантичними зв'язками слів. Найчастіше у список потрапляють синоніми та близькі за тематикою слова.В реальних текстових даних важко слідкувати за появою конкретних слів, маючи лише початкову форму слова. Для розв'язання цієї проблеми створено функцію, яка користується словниковими базами системи. Система за словом може відшукати всі можливі форми слова (парадигми слова).Ці функції надає бібліотека лінгвістичних функцій WordNetUtil.dll, що використовує реалізацію українського та російського WordNet, яка розроблена на кафедрі математичної інформатики, факультету кібернетики Київського національного університету імені Тараса Шевченка.

Робота з системою «LingvoSniffer»

Об'єднує операції, які керують функціями перехоплення мережних пакетів (рис. 1) та збору лінгвістичної статистики (рис. 2).

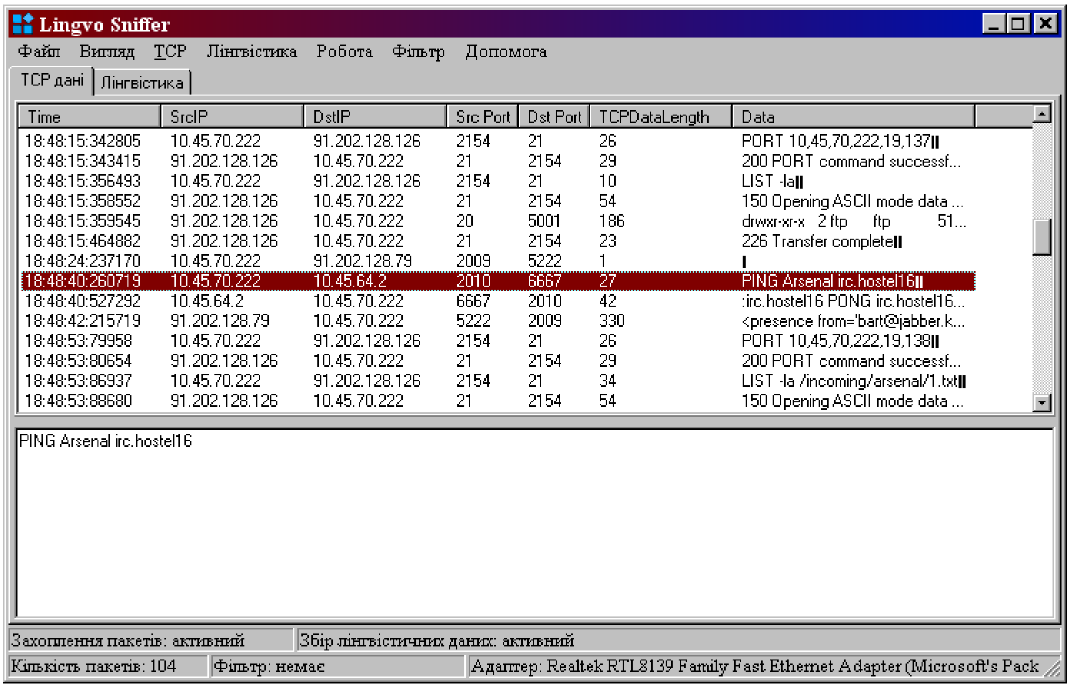


Рисунок 1 – Загальний вигляд системи «LingvoSniffer» та перехоплення мережних пакетів

Доступні наступні функції:
– почати захоплення пакетів – починає захоплення мережних пакетів;
– припинити захоплення пакетів – припиняє захоплення мережних пакетів;
– почати збір лінгвістичних даних – починає збирати лінгвістичну статистику під час захоплення мережних пакетів;
– припинити збір лінгвістичних даних – зупиняє збір лінгвістичної статистики;
– робота лінгвістичної статистики синхронно із захопленням пакетів – вказує, що збір лінгвістичної статистики буде активізовуватися і зупинятися одночасно із активізацією та зупинкою захоплення мережних пакетів;
– конфігурація – змінює конфігурацію програми, пов’язану із словниками та лінгвістичною утилітою.

Найбільш цікавим у даній роботі є лінгвістичний модуль, який допомагає збирати різноманітні статистичні дані. Розглянемо його роботу більш детально.

Робота з лінгвістичною статистикою

Лінгвістична статистика відображається на закладці «Лінгвістика» головного вікна програми. Ця статистика збирається на основі пакетів даних, що були перехоплені (рис. 1).

Статистика складається зі статистичних елементів. Кожний статистичний елемент зберігає такі дані:
– Source IP – IP адреса джерела;
– Source Port – порт джерела;
– Destination IP – IP призначення;
– Destination Port – порт призначення;
– Лічильник – лічильник входження слів;
– Порогове значення – максимально допустиме значення лічильника;
– Стан – вказує на те, чи заблоковане правило;
– Група слів – слова, які шукаються у переданих даних.

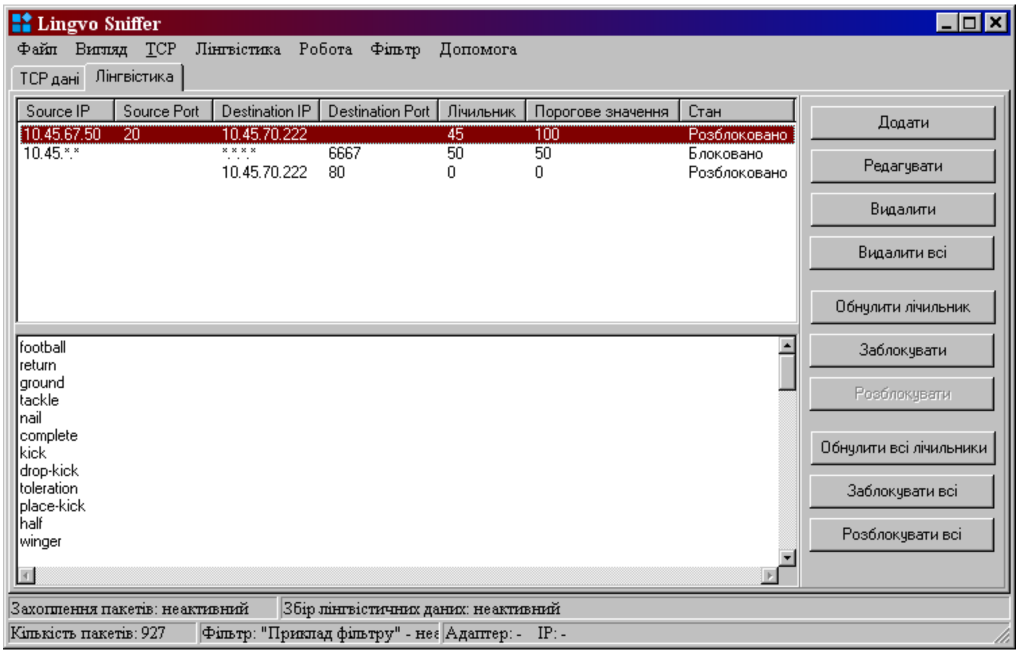


Рисунок 2 – Робота з лінгвістичною статистикою «LingvoSniffer»

Всі дані, крім слів, відображаються в рядках, а слова – в нижньому полі для виділеного елемента. Для збереження статистичних даних потрібно викликати меню «Лінгвістика»->«Зберегти у файл». У діалоговому вікні вказати ім'я файла. Для завантаження статистичних даних з файла потрібно викликати меню «Лінгвістика»->«Завантажити з файла». У діалоговому вікні вказати ім'я файла. Лічильник збільшується, якщо пакет за вказаними критеріями (Source IP, Source Port, Destination IP, Destination Port) містить дані, у яких зустрічаються слова зі списку. Автоматичне блокування в мережі відбувається після досягнення лічильником порогових значень. Якщо порогове значення рівне 0, то автоматичне блокування не відбувається. Блокування відбувається за значеннями Source IP, Source Port, Destination IP, Destination Port, але вже не враховує дані. Додавання статистичного елемента можливе за допомогою кнопки «Додати», що розташована праворуч. Редагування виділеного статистичного елемента можливе за допомогою кнопки «Редагувати», що розташована праворуч. Видалення виділеного статистичного елемента можливе за допомогою кнопки «Видалити», що розташована праворуч. Для статистичного елемента можна обнулити його лічильник, якщо з деякого моменту часу потрібно не враховувати попереднє значення лічильника. Для цього призначена кнопка «Обнулити лічильник». За допомогою кнопки «Заблокувати» можливе ручне блокування елемента. За допомогою кнопки «Розблокувати» можливе ручне розблокування елемента. Кнопка «Обнулити всі лічильники» обнуляє лічильники у всіх елементах. Кнопка «Заблокувати всі» дозволяє вручну заблокувати всі елементи. Кнопка «Розблокувати всі» дозволяє вручну розблокувати всі елементи.

Форма (рис. 3) містить 5 полів для редагування:

– Source IP – IP адреса джерела;
– Source Port – порт джерела;
– Destination IP – IP призначення;
– Destination Port – порт призначення;
– Порогове значення – максимально допустиме значення лічильника, це ціле невід'ємне число.

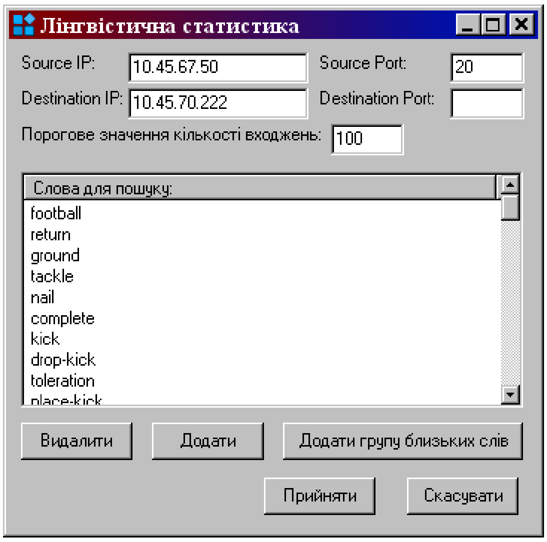


Рисунок 3 – Форма «Лінгвістична статистика»

Також на формі можна редагувати список слів для пошуку. Редагування власне слів можна робити, клацнувши мишею на слові. Для видалення виділених слів призначена кнопка «Видалити». Для додавання одного слова потрібно використовувати кнопку «Додати». Список слів не може бути порожнім. Користувач повинен прийняти зміни або скасувати.

Важливим модулем програми є генерування близьких слів на основі базового слова з використанням бази WordNet. А також генерування парадигми слова. Це дозволить відстежувати синоніми, слова, близькі за тематикою, та визначати всі можливі форми слова.

Для додавання групи слів потрібно натиснути на кнопку «Додати групу близьких слів».

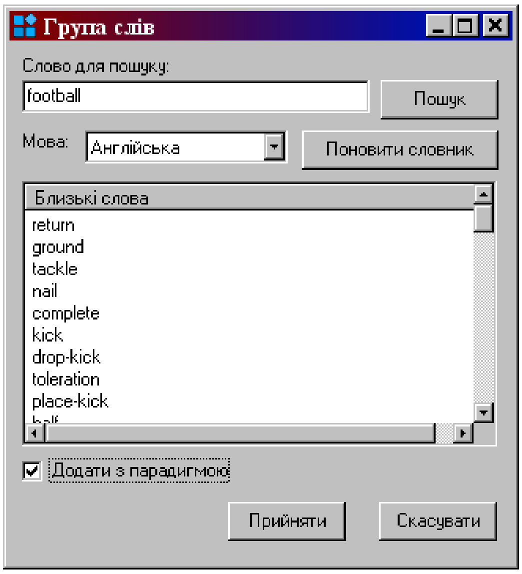


Рисунок 4– Форма «Група слів»

На формі групи слів (рис. 4) ввести початкове слово для пошуку, обрати мову та натиснути кнопку «Пошук».

Доступні мови для обробки:
– англійська;
– російська;
– українська.

Користувач може прийняти список слів чи скасувати.

Висновок

Було проведено дослідження існуючих систем аналізу та захисту мережного трафіка, виявлено їх переваги та недоліки, внаслідок чого визначено вимоги до надійної та зручної системи.

Результатом роботи є реалізація системи перехоплення пакетів напряму з мережного середовища та їх подальшого аналізу. Система робить вторинний аналіз на основі лінгвістичної інформації, що була перехоплена. Є можливість автоматичного та ручного блокування недозволеного трафіка. Програма має гнучке налаштування для забезпечення різноманітних вимог системних адміністраторів та користувачів мереж.

Представлена програмна реалізація системи «LingvoSniffer» може служити базовою для побудови спеціалізованих засобів інших класів, наприклад:
– мережних та протокольних аналізаторів;
– програм моніторингу мережі;
– засобів збору мережного трафіка;
– систем виявлення обміну недозволеної інформації.

Література

1. Топ-рейтинг лучших антивирусов, файерволлов, антишпионов. – Режим доступа: http://www.izcity.com/data/security/article1210.htm
2. Медведовский И.Д., Семьянов П.В., Платонов В.В. Атака через Интернет. – СПб.: «Мир и семья-95», 1997.
3. Протоколы информационно-вычислительных сетей / Под ред. И.А. Мизина и А.П. Кулешова. – М.: Радио и связь, 1990.
4. Halsall F. Data communications, computer networks and open systems. – Addison-Wesley publishing company, 1992.
5. Соломон Д., Русинович М. Внутреннее устройство Microsoft Windows 2000. Мастер класс. – СПб.: Питер; «Русская редакция», 2001.
6. WinPcap documentation. – Режим доступа: http://www.winpcap.org/
7. Архитектура захвата пакетов для Windows WinPCAP. – Режим доступа: http://www.cherepovets-city.ru/insecure/reading/papers/libpcap.htm
8. Lexical database for the English language. – Режим доступа: http://wordnet.princeton.edu/

А.В. Деревянченко

Анализ информации в компьютерной сети с помощью системы «LingvoSniffer»

В статье рассматриваются принципы построения системы анализа информации в компьютерной сети с применением лингвистического модуля, которая может решать актуальные вопросы контроля сети как средства передачи и обмена данными между пользователями и доступа к определенной информации.

Стаття надійшла до редакції 21.07.2008.